(19) Europäisch s Patentamt
European Patent Offic
Office européen des brevets

(11) EP 0 957 648 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
17.11.1999 Bulletin 1999/46

(51) Int Cl.[6]: **H04Q 3/66**, H04Q 3/00

(21) Application number: **99303448.7**

(22) Date of filing: **30.04.1999**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **30.04.1998 CA 236427**

(71) Applicant: **NORTEL NETWORKS CORPORATION**
Montreal Quebec H2Y 3Y4 (CA)

(72) Inventors:
- **Beshai, Maged E.**
  Stittville, Ontario K2S 1p2 (CA)
- **Vickers, Richard**
  Kanata, Ontario K2K 1P2 (CA)

(74) Representative: **Bewley, Ewan Stuart**
Nortel Networks,
IP Law Group,
London Road
Harlow, Essex CM17 9NA (GB)

(54) Method and apparatus for distributed control of a multi-class network

(57) A method and apparatus for control of a multi-class digital network (20) are described. The network supports a plurality of digital services using dynamically-configured service bands (36) to support various transport modes and qualities of service. Flexible controls with two degrees of freedom are used at the node level to dynamically configure the service bands. The first degree of freedom is realized using adaptive alternate routing to balance the traffic intensity across the network and, hence, increase the efficiency of the bands. The second degree of freedom is achieved through adaptive network partitioning to provide elasticity to the band structure. This permits band capacity to be adaptively reconfigured as connections are added and dropped. The switching nodes (22) use rate controllers (50) to divide the capacity of each link (24) among the connection classes according to rules which ensure consistent service attributes for each band across the network. The rate controllers are adapted to support different transport modes and different packet sizes so that the deconstruction and reconstruction of packets at network edge devices is eliminated. The advantage is a flexible multi-class network which dynamically reconfigures to service traffic loads as service demands fluctuate.

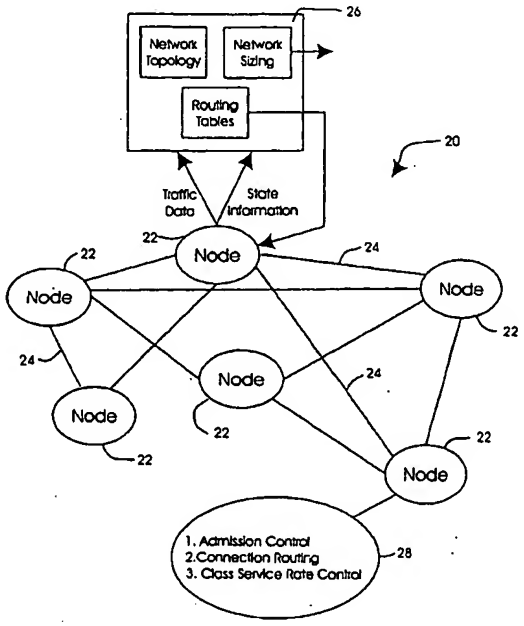


FIG. 1

EP 0 957 648 A2

D scription

TECHNICAL FIELD

[0001] This invention relates generally to multi-class digital networks and, in particular, to methods and apparatus for the distributed control of a multi-class digital network in order to efficiently adapt to rapid fluctuations in traffic load and service demand.

BACKGROUND OF THE INVENTION

[0002] A multi-class digital network must accommodate traffic having different characteristics, quality of service (QOS) requirements, and transport modes. For example, a multi-class backbone network may be required to support both connection-based traffic and connectionless traffic. Further complication is introduced by the fact that within each category the traffic may include several classes differentiated by their characteristics and service requirements. The traffic classes may require different and possibly conflicting controls in order to satisfy service commitments and ensure acceptable service quality.

[0003] Asynchronous Transfer Mode (ATM) switching technology was developed to provide multi-service digital transport. It was assumed that ATM networks would provide the flexibility and quality of service required to satisfy the demand for digital services.

[0004] One disadvantage of ATM networks is that they are adapted to transport packets of only one size and format. Consequently, the packets of services which do not use ATM format must be deconstructed on admission to the ATM network by network edge devices and reconstructed by network edge devices on egress from the ATM network. This slows service delivery, increases computational requirements and complicates the structure and functionality of edge device interfaces.

[0005] The varying service requirements in a multi-class network are difficult to satisfy without traffic segregation and network partitioning. For example, certain services such as voice and video are somewhat loss tolerant but delay intolerant, while other services such as the exchange of data packets between computers are quite delay tolerant but completely loss intolerant. In accommodating such variations in service, a multi-service network naturally segregates into a plurality of layers or "bands" which respectively serve the requirements of different types of traffic. This natural division of a network into service bands is well understood and has been widely discussed in the relevant literature.

[0006] A challenge in network management is designing network routing and admission controls to manage the service bands in a multi-class network to efficiently accommodate fluctuating service demands. It is well understood that while total network traffic may change relatively slowly over time, the traffic mix in a multi-class network may fluctuate rapidly and unpredictably. To date, efficient methods of accommodating rapid and unpredictable fluctuations in traffic service demand have eluded network designers and traffic managers. There therefore exists a need for a multi-service, multi-class digital network and methods for controlling the network which can accommodate the increasing demand for digital services without unreasonable investment in network infrastructure.

OBJECTS OF THE INVENTION

[0007] It is an object of the invention to overcome disadvantages with existing methods systems.

[0008] It is a further object of the invention to provide a method of controlling traffic admission and routing in a multi-class digital network wherein admission control is performed by computing an equivalent bit rate for each network admission request at a node in the network and traffic admission is based on the equivalent bit rate.

[0009] It is yet a further object of the invention to provide a novel method of computing an equivalent bit rate for a traffic admission request which significantly reduces computational effort in nodes subsequent to an edge node which receives the request.

[0010] It is yet a further object of the invention to provide a method of controlling traffic admission and routing in a multi-class digital network which supports a plurality of routing schemes.

[0011] It is yet a further object of the invention to provide a method of controlling traffic admission and routing in a multi-class digital network in which routing schemes are linked to classes of service.

[0012] It is yet a further object of the invention to provide a link controller for a transport link in a multi-class digital network which stores a sampling frequency for each class of service being served at any given time by the link controller to ensure that the class of service is guaranteed a minimum service rate.

[0013] It is another object of the invention to provide a method of distributed control of a multi-class digital network in which traffic normally allocated to a path is overflowed to a connection if the path is filled to capacity.

[0014] It is a further object of the invention to provide a method of distributed control of a multi-class digital network in which traffic belonging to a class that is normally served by connection is served by a path if the path is temporarily idle.

[0015] It is yet another object of the invention to provide a method of distributed control of a multi-class digital network in which connectionless traffic without a quality of service guarantee is transported in time slots not used by other traffic having a guaranteed quality of service.

[0016] The above objects are met by the combinations of the features of the main claims, the sub-claims disclose further advantageous embodiments of the invention.

SUMMARY OF THE INVENTION

[0017] The summary of the invention does not necessarily disclose all the features essential for defining the invention; the invention may reside in a combination of the sub-features of the invention.

[0018] In accordance with a first aspect of the invention there is provided a method of controlling traffic admission and routing in a multi-class digital network serving variable-size packets, by computing an equivalent bit rate for each traffic admission request as received at a node in the network, determining whether a connection for the traffic admission request can be established through the network, and establishing the connection if a connection can be established, **CHARACTERIZED BY**:

adding the equivalent bit rate for the traffic admission request to a current service rate allocation of a class of traffic being served by an egress link through which the connection is established, to permit a service rate controller to control transmission of each class of traffic on the egress link.

[0019] The invention further provides apparatus for distributed control of a multi-class network in which network processing is distributed between a network control element and network nodes to mimimize computational effort and maximize transmission efficiency, **CHARACTERIZED** in that:

the network control element receives traffic state and network state information form the nodesand uses the information to compute network traffic routing sets that are distributed to the nodes;
the network nodes include node control elements that control traffic admission, traffic routing and the computation of service rate allocations for classes of traffic served by egrees links at the node; and
link controllers for the links, the link controllers including service rate controllers adapted to control an egress of variable-sized packets on the link, at a guaranteed service rate.

[0020] The multi-class digital network in accordance with the invention distributes the network processing load between a network control element which handles global functions that are best performed at the network level and traffic functions which are best performed in a distributed fashion at the node level of the network. This distribution of functionality minimizes computational effort and maximizes transmission efficiency.

[0021] The network control element in accordance with the invention receives traffic intensity and network state information from the nodes in the network which periodically report such information to the network control element. Using the network state information, the network control element maintains a network topology. The network topology and the traffic intensity data are used by the network control element to compute network traffic routing sets which are identified to the nodes along with an order of preference. The computed routing sets are distributed to the network nodes and used by the network nodes in processing traffic admission requests.

[0022] The network nodes include node control elements which control traffic admission, traffic routing and the computation of service-rate allocations for classes of traffic served by egress links at the node.

[0023] Edge network node control elements receive traffic admission requests from subtending sources. The edge node control elements compute an equivalent bit rate for each traffic admission request based on a novel method in accordance with the invention. In order to minimize further processing when it is necessary to establish a connection across the network, the edge node control element also computes variables which enable subsequent nodes involved in the connection to rapidly compute an approximate equivalent bit rate used in route selection.

[0024] The multi-class digital network in accordance with the invention preferably supports a plurality of digital services which may require different transport modes, each transport mode consisting of at least one transport protocol. Since the packets of the same protocol may be of variable size, a link controller is provided which accommodates variable packet sizes so that packets need not be disassembled and converted to a standard format by edge device interfaces.

[0025] The multi-class digital network adopts a service-rate discipline comprising a guaranteed minimum rate per class in order to ensure efficient use of the network while meeting transmission rate and quality of service commitments. In order to minimize computing requirements for routing, high-frequency, low bit-rate traffic is preferably served by paths commonly referred to as direct routes set up through the network. High bit-rate connection-oriented traffic is preferably served by connections set up on demand. Each service type is preferably assigned to at least one class and each class is preferably assigned to a separate band in the network. The bands are dynamically configured and have elastic boundaries which fluctuate with traffic load. Unused time slots accept traffic from any waiting source in a

predetermined hierarchical order in which connectionless traffic without a quality of service is served last.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] The invention will now be further explained by way of example only and with reference to the following drawings wherein:

Fig. 1 is a schematic diagram of a multi-class digital network in accordance with the invention;
Fig. 2 is a schematic diagram showing cost ellipses used by a network control element to compute routing sets for the network node shown in Fig. 1;
Fig. 3 is a schematic diagram of an example network used to illustrate determining the routing sets for network nodes by a network control element;
Fig. 4 is a routing table for a destination node 11 shown in Fig. 3;
Fig. 5 is a schematic diagram showing network links partitioned into banks corresponding to classes of traffic;
Fig. 6 is a schematic diagram illustrating network paths and network connections through a sample network;
Fig. 7 is a graph illustrating the relationship between transport efficiency and processing throughput in relation to a network which uses connections or paths;
Fig. 8 is a schematic diagram depicting intraband and interband controls in a multi-class digital network in accordance with the invention;
Fig. 9 is a schematic diagram illustrating capacity management in a multi-class digital network with paths and connections where connection-level sharing is practised;
Fig. 10a is a schematic diagram illustrating capacity management with paths and connections in a network link;
Fig. 10b is a schematic diagram illustrating capacity management, as in Fig. 10a, except that a connection band is not subdivided, and any unallocated capacity is granted to connectionless-mode traffic;
Fig. 11 is a schematic diagram illustrating a rate controller in a node control element for a transport link in a multi-class digital network in accordance with the invention;
Fig. 12 is a schematic diagram of class allocations for variable packet sizes in a multi-class, multi-service network in accordance with the invention;
Fig. 13 is a schematic diagram illustrating a preferred arrangement for processing connection admission requests at node control elements in accordance with the invention; and
Fig. 14 is a graph illustrating a method used for minimizing processing required to compute equivalent bit rates as connections are established across a multi-class digital network in accordance with the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0027] Fig. 1 is a schematic diagram of a multi-class network in accordance with the invention, generally indicated by the reference 20. The network includes a plurality of switching nodes 22, hereinafter referred to simply as "nodes". The nodes are interconnected by transport links 24, hereinafter referred to simply as "links". The links 24 may be any one of a number of types of transport links well known in the art. The links 24 are designed to support a transfer rate appropriate for the traffic loads that they must transport. The multi-class network 20 in accordance with the invention includes a network control element 26 which preferably performs at least the functions of network topology monitoring, and computing and distributing network traffic routing sets to the nodes 22, as required. The network control element 26 also preferably performs the function of network sizing computations which produce periodic specifications for the sizes of the inter-nodal links 24. The network control element 26 is, for example, a server connected to a one of the nodes 22, or the like.

[0028] In the multi-class network 20, other network functions are preferably distributed among node control elements 28. Consequently, the node control elements 28 perform traffic admission control, connection routing for connection-oriented traffic and the computation of service-rate allocations for classes served by egress links managed by a node control element 28. Each of these functions will be described below in detail. Even though only one node control element 28 is shown in Fig. 1, it will be understood that each node 22 includes a node control element 28

[0029] The multi-class network 20 may carry classes of traffic having distinctly different characteristics and service requirements. The different service requirements for the various traffic classes are difficult to meet without traffic segregation and network partitioning into layers or "bands". For example, open-loop controls and closed-loop controls have conflicting attributes and require distinctly different handling. Both can, however, operate efficiently in the same network if each control is applied to a separate network band. Connection-based traffic and connectionless traffic also have other distinctly different service requirements. For example, human-to-human communications such as voice and video are somewhat loss-tolerant in that a certain number of packets can be lost without detection by the human participants, yet they are quite delay-intolerant in that any delay or "jitter" in packet delivery is quite noticeable. Data

communications between computers on the other hand are quite delay-tolerant but strictly intolerant to packet loss.

[0030] Network segregation into bands is preferably done so that each network band corresponds to a traffic class. The same control method applied to a given class should be practised across the network. In order to efficiently use the network, the boundaries between successive bands must be elastic to permit band capacity to expand and shrink in response to variations in network traffic.

[0031] Topological design and overall network dimensioning are based on long-term traffic forecasts and other considerations well known in the art. Aggregate network design is a slow process which takes place over time. Network sizing is therefore preferably a function performed by the network control element 26. Network sizing requires traffic characterization. Performance estimation is a function of traffic load and estimates of spatial traffic distribution. Given the difficulty of traffic characterization in a multi-band, multi-service network, effective network sizing computations may be based on traffic and performance measurements. As shown in Fig. 1, the network control element 26 therefore receives traffic measurement data from the node control elements 28. The traffic measurement data is transferred to the network control element 26 through the network. The traffic measurement data is accumulated in appropriate tables and periodically analyzed in order to determine appropriate sizes for the links 24 in the network. The entire network 20 must be provisioned to serve the combined traffic load of all classes. The combined traffic load is, of course, variable over time, however, the combined traffic load is less volatile than the traffic loads of the individual classes. The network traffic data may be sorted to determine traffic loads on each inter-nodal link 24 and the data can be used to determine when a link size is inadequate or when a direct link between unlinked nodes is warranted. Consequently, in addition to traffic data each node control element 28 involved in traffic admission control should report to the network control element 26 the details of admission request denied and the cause of admission failure. The frequency at which an analysis of this data is conducted depends on a number of factors well known in the art, including storage capacity at the network control element 26.

[0032] The node control elements 28 also pass network state information messages to the network control element 26 to permit the network control element 26 to maintain topological awareness of the network 20 and its state. In some networks, a topological map of the network is maintained in every node. Consequently, "flooding" of state information is required. This flooding of information uses network resources and ties up node computing cycles. In order to free resources for transport control, in the multi-class network in accordance with the invention, a complete topological network map is maintained only at the network control element 26. To ensure that the topological information is accurate and complete, each link 24 is preferably monitored by each of the nodes 22 at its opposite ends. Consequently, if the status of a link 24 changes, the change is reported by two nodes 22 to the network control element 26. Likewise, the state of each node 22 is monitored by two neighbouring nodes 22 appointed by the network control element 26. If any node 22 malfunctions or becomes inoperable, the network control element 26 appoints alternate monitor(s) for any other nodes that were monitored by that node. This method ensures that the network control element 26 receives accurate network state information to guarantee that network topology is accurate and up to date.

[0033] The network topology is used, as required, to compute network routing sets which are used by node control elements 28 to route traffic admission requests across the network. Direct routes, which may include direct links or shortest route paths through the network, are the routes of first preference. Indirect routes from an originating node 22 to a destination node 22 are preferably sorted according to cost ellipses 30 illustrated in Fig. 2. In using the cost ellipses 30 to sort route sets, the origin and destination nodes, X and Y, are placed at the two foci and paths along intermediate nodes are sorted according to cost as determined by an acceptable model. Transport cost models are well known in the art and a number of good models are publicly available. Any path having intermediate nodes which fall between the same elliptical cost boundaries is considered to be of comparable cost. For example, the paths x-d-y and x-f-y in Fig. 2 are of comparable cost and are considered as equal alternatives.

[0034] Fig. 3 is a schematic diagram of an example network for illustrating the route sets computed by the network control element 26. Each node control element 28 is provided with a list of candidate routes to each destination. The list from node 0 to node 11 includes a direct route, a set of first-choice routes and a set of second-choice routes, etc. The list includes only the first adjacent node along the shortest path to the destination. Routes from each origin to the destination are determined according to cost criteria described above. The routes are sorted according to static cost and assigned a selection priority. Each network node control element 28 maintains a table of routes received from the network control element 26. Whenever network topology changes, the network control element 26 determines which links 24 are affected by the change and recomputes network routing sets accordingly. In order to conserve time and ensure that nodes use effective routing sets, only those links 24 affected by a change in network topology are recomputed. When a new route set is recomputed for any particular network node 22, the recomputed route set is preferably transferred immediately to the associated node control elements 28.

[0035] Fig. 4 shows an exemplary routing table computed by th network control element 26 for the destination 11 from each of the other 10 nodes in the network. As is apparent, a direct route is specified whenever it is available. The direct route is always preferred in traffic routing if free capacity exists on that route. As an alternative to the direct route, a first set of lowest-cost routes is specified. For example, from node 0 to node 11, the first alternate set includes nodes

5 and 6. The network control element 26 specifies only the immediately adjacent node for route selection. As will be explained below, during traffic admission the node control elements 28 select the two immediate links which are least occupied to find a route to service an admission request.

[0036] In a multi-class network in accordance with the invention, network traffic is divided into a manageable number of classes, whose definition preferably depends on:

1) the transaction format (synchronous transfer mode (STM), asynchronous transfer mode (ATM), Internet protocol (IP), etc.);
2) the method of transaction processing, e.g., connection-oriented or connectionless; or
3) the method of flow control (open-loop or closed-loop).

The term "transaction" as used in this document is to denote the process of transferring information from a source to a sink in a single session. A class may include several connections of similar traffic or may represent a single connection which has a high bit-rate requirement. A class may also include connectionless traffic with no quality of service (QOS) requirements.

[0037] Fig. 5 shows the links 24 of a network node 22 partitioned into a plurality of bands, each of which supports one class of traffic. The total capacity of each link 24 is divided among the classes it supports. The links 24 may be partitioned differently since the same mix of traffic is not necessarily transported over each link. Different routing schemes and admission criteria may be applied to the various network bands. For example, a voice service class may use a simple routing scheme while classes of higher-speed connections may use an elaborate state-dependent selective routing scheme as will be explained below in some detail. In addition to the difference in routing and admission criteria, operational measurements, tariff setting and billing procedures may differ among the classes of traffic.

[0038] In order to maximize transport efficiency and minimize processing load, at least certain classes of traffic preferably use paths through the network for traffic routing. Fig. 6 schematically illustrates a relation between a path and a connection. In Fig. 6, a solid line through a node indicates a path 32 between non-adjacent nodes and the line connecting nodes indicates a connection 34. A path and a connection may exist on the same link at the same time and within the same class. A path 32 is a reserved route between two nodes 22, possibly traversing several intermediate nodes 22. Intermediate nodes 22 in a path are involved in the path allocation process only when the path is first established or when its capacity allocation is modified. A path 32 may accommodate a large number of connections, possibly several hundreds, between its origin and destination nodes. The originating node 22 views the path as a direct connection to the destination node 22. The intermediate nodes 22 in a path are not involved in the admission of the individual connections belonging to the path 32. A path 32 may serve more than one class of traffic. However, the class definition is only relevant to the end nodes 22.

[0039] A connection 34 seeks admission at a specific ingress point in the network. A connection specifies a destination, but there is no fixed route between the originating node and the destination node of the connectiion. Rather, a route is negotiated when the connection is set up. Each node 22 traversed by a connection is involved in the interrelated decision regarding the admission, or otherwise of the traffic admission request for a connection and the actual selection of the end-to-end route for the request.

[0040] Paths and connections may or may not have a capacity reservation. A capacity reservation is necessary to achieve a specified QOS. Without a capacity reservation, the QOS is based only on a weighted priority, or some other class differentiation, and the path 32 or connection 34 are used only to facilitate the forwarding of packets at the nodes.

[0041] As shown in Fig. 7, paths maximize processing throughput since admission control may be handled exclusively by an edge node which receives a traffic admission request. Establishing paths, however, is not the most efficient use of network resources unless adequate stable traffic exists to keep the path full. In general as shown in Fig. 7, transport efficiency decreases as the paths' allotment in the network increases. Conversely, transport efficiency increases with the increase in the connections' allotment, however processing throughput is significantly affected.

[0042] In the network 20 in accordance with the invention, flexible controls which permit two degrees of freedom are realized by employing adaptive routing in addition to adaptive network partitioning. The two degrees of freedom are depicted in Fig. 8 which illustrates the intra-band and inter-band controls that may be used in a network 20 in accordance with the invention. Within each network band 36, adaptive alternate routing may be used to balance the traffic intensity across the network and thereby increase the efficiency of the band. The elastic boundaries of a band 36 may vary slowly, typically in seconds, between successive changes while adaptive routing is applied on a per-connection basis. The two degrees of freedom compliment each other to ensure a balanced, efficient network. The principal control element within a band is the routing scheme. State-dependent routing results in load distribution across the network band. The load distribution is further improved with selective state-dependent routing.

[0043] As will be explained below, the network nodes 22 use rate controllers to divide the capacity of each link among the network bands 36 in accordance with predefined rules. Because a large proportion of connections traverses more than one link 24 in the network 20, the rate allocations for the bands cannot be done independently at each node. The

rate allocations must be coordinated amongst the nodes in order to ensure that the physical constraints of the link capacity are observed and that end-to-end service requirements are met.

[0044] As described above, a n twork band 36 may include both paths 32 and connections 34. Low bit-rate admission requests which occur at high frequency are prime candidates for paths 32. High bit-rate connections which are requested at a low frequency are better served by independent connections 34. By appropriately selecting path sizes, and with proper splitting of traffic between paths 32 and connections 34, a network divided into bands 36 with elastic boundaries can be almost as bandwidth-efficient as a fully shared network, while call processing load is reduced and QOS differentiation is facilitated. The word "bandwidth" as used in this document is intended to denote capacity in bits/second.

[0045] The prior art approach to configuring a communications network has been to seek an optimal tradeoff between transport efficiency and processing efficiency. Consequently, traffic was divided between a set of paths 32 and connections 34. Paths 32 consume less processing resources than connections 34. However, due to the random fluctuations of traffic intensity, a path 32 may occasionally suffer from low utilization. This is particularly the case for low-intensity traffic streams which are normally quite variable in their volume. Connections 34 require more processing capabilities due to the signaling load and processing at one or more intermediate nodes 22.

[0046] The traffic efficiency of a path 32 can be increased by allowing limited queuing of traffic admission requests at edge nodes. Normally, if a path 32 does not have sufficient free capacity to accommodate a traffic admission request, the traffic admission request is either overflowed to a connection band or it is refused. However, permitting traffic admission requests to wait until a sufficient free capacity in the designated path becomes available, or a time-out threshold is reached, significantly improves the path utilization while decreasing overall network processing effort. A compromise arrangement in accordance with the invention is to use paths 32 to reduce processing effort combined with a shared capacity allocation for connections 34. Such an arrangement is shown in Fig. 9. Traffic admission requests that occur too infrequently to justify establishing a path 32 may use the connections allocation. As shown in Fig. 9, several paths are permitted to overflow to the connection allocation which may be within the same network band 36. Each of the paths shown in Fig. 9 is sized to accommodate most of the traffic between its end nodes. At the connection level, traffic admission requests allocated to a path may overflow to the shared connection allocation if the path cannot accept the admission request. At the data level, packet-type connection traffic may use vacant time slots allocated to paths, as illustrated in Fig. 10a. This significantly increases the traffic capacity of the link and brings the overall bandwidth efficiency closer to that of a fully-shared link without affecting the path's performance. The optimal splitting of a link capacity between paths and connections is determined dynamically to follow the traffic load variation.

[0047] Fig. 10b shows a preferable pattern of partitioning the network transport capacity among the network links. End-to-end paths 32 are used for direct connections, and are established whenever the end-to-end traffic volumes exceed a certain threshold. The remaining capacity in each link, if any, is divided into bands, each of which corresponds to a traffic class. The capacity of each band is dynamically adjusted by the connection-admission process. The capacity of a band increases with each new connection admission, and decreases with each connection departure. Any leftover capacity is used for connectionless-type traffic. The arrows in Fig. 10b indicate that packets belonging to connections of any traffic class may exploit the unused time slots allocated by the rate controller to any path. The reverse is not allowed; packets belonging to a path may not be transmitted during unused time slots allocated by the rate controller to any connection class.

[0048] In order to ensure that each traffic class is guaranteed a minimum service rate, it is necessary to provide some form of service-rate controller at the link level. In accordance with the invention there is provided a link service-rate controller shown in Fig. 11 and generally referred to by the reference 50. The combined service-rate allocations for all classes of traffic being served by an egress link must satisfy the condition:

$$0 < f_j = \frac{F_j}{R} < 1, j = 0,\ldots,K-1, \quad with \quad \sum_{j=0}^{K-1} f_j < 1$$

wherein:

K is the number of classes;
R is the link rate in bits per second; and
F_j is the required service allocation for class j in bits per second; and
f_j is the normalized service rate for class j.

[0049] Since the allocated service rate per class is less than the link rate, and since it is not possible to transfer a fraction of a packet at any time, it is necessary to wait for several clock cycles to be able to transfer a first packet from

a given class buffer. In the link s rvice-rate controller shown in Fig. 11, which shows four classes (K=4), a clock pulse transferred on a clock line 52 provides a clock signal to sampling frequency circuits 54. The clock pulse on clock line 52 is regulated to a predetermined nominal packet size, preferably equal to the minimum sized packet transfer rate of a link 24 served by the link service-rate controller 50. Other clock speeds may also be used. Consequently, the clock pulse can be set to a rate which is some integer multiple of the minimum size packet transfer rate.

[0050] Each sampling frequency circuit 54 includes a memory register which stores a class service allocation represented by the character "Δ". This is, hereafter, called the class credit. The value of Δ is dynamically computed for each of classes 0 through K-1. The class service allocation Δ is preferably a floating point number between 0 and .99. When a connection admission request is accepted into a class band, an equivalent bit rate (EBR) normalized to the link capacity is added to the value of Δ. When the connection is released, the normalized EBR is subtracted from Δ. The process for computing EBR is described below in detail.

[0051] The credit Δ for a given class is the ratio of the required service rate of the class (occasionally called the "bandwidth of the class") divided by the capacity of the link under consideration. For example, the credit Δ of a class requiring a 100 Mb/sec in a link of capacity 620 Mb/sec is approximately:

$$\Delta \equiv 0.16.$$

Note that Δ is dimensionless.

[0052] The accumulated credit for a given class is to be compared with the normalized packet size. The normalized packet size is the packet length (in bytes, say) divided by a nominal packet size (64 bytes, say) which is applied uniformly for all classes traversing a link. Preferably, the normalized packet size should be standardized across the network using a value of 64 bytes, for example.

[0053] Each time a clock signal is received by the sampling frequency circuit 54, the value of Δ stored in the memory register is added by a adder 56 to a memory register which accumulates a class allocation sum. A comparator 58 compares the class allocation sum with the "normalized packet size" 64. If a class buffer 62 is empty, the class allocation sum is set to zero.

[0054] Fig. 12 shows a schematic representation of the transmission of packets of variable size. At a top of the figure, a schematic representation of clock cycles is shown. At each clock cycle δ, which equals the transfer time for a predetermined nominal packet size, the value Δ is transferred from the sampling frequency circuit 54 (Fig. 11) to the adder 56. At each clock signal, the comparator 58 compares the value of the memory register in adder 56 with the normalized packet size 64.

[0055] The accumulated credit in the adder 56 is stored in the memory register of the corresponding comparator 58 and used for comparison with the contents of adder 56 memory register. When the accumulated sum in the adder 56 is greater than or equal to the normalized packet size 64, a selector 66 writes an eligible class number 68 in the ready queue 60 and the comparator 58 subtracts the normalized packet size 64 from the accumulated sum in the adder 56 so that the balance is added to the class allocation Δ at the next clock signal, as shown in the table in Fig. 12. When an eligible class number 68 arrives at the head of the ready queue 60, the link service-rate controller 50 selects a packet at a head of the corresponding class buffer 62 and transmits that packet over the link.

[0056] If the ready queue 60 becomes empty because there are no classes with a specified QOS to be transferred, traffic without a specified QOS is transferred until another class number 68 appears in the ready queue. In order to ensure equitable treatment, a rotating pointer is preferably used to track a next class to be served when the ready queue 60 is empty.

[0057] The traffic transferred through the network is preferably classified by the node control elements 28 into three types: connection-based traffic with a specified QOS; connection-based traffic without a specified QOS; and connectionless traffic without a specified QOS. Each node control element preferably monitors the amount of the connection-based traffic without a specified QOS and assigns it an appropriate service rate if there is an appreciable amount of that type of traffic and there is unused capacity on the link. If the link capacity is required by traffic with a specified QOS, however, the service rate is withdrawn from that traffic.

[0058] As noted above, an important aspect of the invention is the distributed control of routing through the network. The network in accordance with the invention preferably supports a plurality of routing disciplines. The preferred routing disciplines are:

1) shortest path hop-by-hop routing;
2) selective routing by a conservative scheme; and
3) selective routing by a true state scheme.

[0059] Preferably, the different routing disciplines are assigned to different classes in order to best serve the require-

ments of the class.

[0060] In the shortest path hop-by-hop routing, each node has a list of candidate routes to a destination. The list includes a direct route, if one exists, a set of first-choice routes, and a set of second-choice routes, etc. As described above, the list includes only the next node 22 to a destination along the shortest path. The routes from origin to each destination are determined according to cost criteria, as also described above. Each node only maintains information about the occupancy of its outgoing links. When a new traffic admission request is received at a node 22, the node control element 28 first checks the occupancy of its outgoing link associated with its direct route, if one exists. If the direct route is full or not available, the node control element 28 compares the occupancy state of the outgoing links of the first-choice set of routes and selects the most vacant route. If the available capacity is still not sufficient to accommodate the new traffic admission request, the second-choice set of routes is inspected, and so on. The number of hops to a destination using a given egress port is known by the node control element 28. The edge node control element 28 therefore assigns a number of "route selection credits" to a traffic admission request which is equal to the number of hops along path selected by the edge node to the destination, plus 2 to allow some variation in downstream route selection to accommodate link congestion. The traffic admission request is then forwarded to the next node selected from the routing table. Upon arrival at the next node, the node control element 28 of that node deducts 1 from the route selection credits assigned by the edge node control element 28, and checks the availability of its link to the shortest path to the destination. Each traversed node 22 subtracts 1 from the route selection credits. By ensuring that the number of remaining hops along the path is not greater that the available route selection credits, and by avoiding a return to the immediately preceding node, the route is guaranteed to be cycle-free. By always looking for the most vacant link in the preferred order, the route is guaranteed to be efficient. Shortest path hop-by-hop routing is best suited for low-bit-rate traffic where strict route optimization is not warranted.

[0061] For high-volume low or medium bit-rate traffic where more powerful route optimization is warranted, a conservative routing scheme may be used to route connections through the network 20 in accordance with the invention. A conservative routing scheme for ATM networks was published by M. Beshai and J. Yan in November of 1996 at the International Network-Planning Symposium in Sydney, Australia in a paper entitled "Conflict-free Selective Routing in an ATM Network". In accordance with the conservative scheme, when a traffic admission request is received at an edge node 22, the node control element 28 computes an EBR for the connection and examines the availability of its outgoing links in the preferred order to determine which link(s) can accommodate the traffic admission request. On locating at most two available links with adequate bandwidth, the node control element 28 deducts the EBR from the available capacity of each link, computes the route selection credits for the traffic admission request message and forwards the traffic admission request message to the next node(s) in the route path(s). Meanwhile, other traffic admission requests permitted to use the conservative routing scheme are allowed to proceed, even though the prior traffic admission request may be denied at some other point in the network and, consequently, the available capacity of the link(s) may not reflect the true state of the link's current usage. Since only relatively low or medium bit-rate connections use this scheme, the waste of network resources is minimal. The purpose of the conservative routing scheme is to find a lowest-cost route for the connection in a reasonable time without sacrificing too many network resources.

[0062] The third routing discipline is a true-state scheme in which traffic admission requests are processed sequentially and, while a true-state connection is being set up, no other traffic admission request is permitted to proceed in the same class of service. A true-state routing scheme for an ATM switching network is described in applicants' United States Patent No. 5,629,930 which issued on May 13, 1997.

[0063] When a traffic admission request is received at an edge node and the traffic admission request is determined to belong to a class of service which requires true-state routing, all other traffic admission requests in that class of service are held until a connection for the request is established or denied. In true-state routing, the node control element 28 searches its links in the preferred order for the two links which have the highest free bandwidth to serve the request after the EBR for the connection request is computed. If a link(s) is found, the route selection credits are computed for the traffic admission request and the request is forwarded to a next node(s) in the route. Thereafter processing proceeds as described above until the connection is established or denied because of a lack of link capacity at some point in the route. While the true-state scheme has the advantage of admission given the true bandwidth capacity of the available link(s), it has the disadvantage of delaying the progress of other connections in the same class. Note that connection admission processing for other classes can proceed while a connection belonging to the true-state class, if any, is being processed.

[0064] It is estimated that in a typical case about 300 conservative scheme traffic admission requests can be processed per second assuming a processing time of about 1 millisecond per request. For a processing true-state routing, it is estimated that only about 20 requests per second may be processed for an average link length of about 200 kilometers.

[0065] Preferably, a network 20 in accordance with the invention uses a hybrid scheme in which true-state routing coexists with conservative routing and shortest path routing. In that instance, true-state routing is limited to connections with high EBR values while conservative routing is used for connections with lower EBR values, in order to increas

the permissible attempt rate. Hop-by-hop routing is used for lowest EBR values where there is no value in using resources to find a least-cost route.

[0066] As described above, the EBR must be computed for each traffic admission request received at an edge node 22 of the network 20. The value of the EBR is determined by traffic descriptors; QOS; and service rate for the class. The traffic descriptors include a peak rate of emission from the source; a mean rate of emission from the source; and, a mean packet size. These values are generally provided by the source when a traffic admission request is received. In some cases, however, they can only be determined by measurement. A QOS request is submitted with the admission request and defines the delay tolerance and loss tolerance for the connection.

[0067] In a single class network, the service rate at an egress port in a node is the entire link rate. In a multi-class, multi-band network where the bands do not share their free time slots the service rate for a class is the capacity allocated to the band associated with the class. Since the service rate fluctuates dynamically, the boundaries between classes are elastic and the capacity per class varies over time. This means that an EBR for a given connection which is computed at a given instant may need to be revised when the band capacity changes. Such recomputation is impractical.

[0068] In the method in accordance with the invention, a policy of guaranteed minimum capacity allocation along with sharing among the bands is adopted. The computation of the EBR for a connection in any band can therefore be based on the entire link capacity, thus eliminating the need to revise EBR calculations. When a traffic admission request is accepted, the computed EBR is added to the memory register in the sampling frequency circuit 54 of the link service-rate controller 50 (Fig. 11). When a session ends and the connection is torn down, the EBR is deducted from the memory register in the sampling frequency circuit 54. Thus the size of a band 36 for a class of traffic fluctuates with traffic load. The EBR is readily calculated using, for example, the extended Gibbens-Hunt formula which is known in the art. The computation is preferably completed in the edge node while new traffic admission requests are queued for treatment, as shown in Fig. 13. If the EBR is calculated while the traffic admission request is queued, traffic admission request delay is reduced. In accordance with the method described above for computing EBR, the EBR is dependent on link capacity. It is therefore necessary when setting up a connection through several nodes 22 to compute an EBR at each intermediate node. In order to accomplish this, prior art methods forwarded the traffic descriptors and QOS along with a traffic admission request message, and each subsequent node involved in the admission request recomputed the EBR using the extended Gibbens-Hunt formula which is known in the art. Other methods can also be used. For example, the EBR can be calculated using the Buffet-Duffield formula, as described in United States Patent Application No. 08/723,649 filed October 3, 1996.

[0069] In accordance with the methods of the invention, recomputation is minimized and connection request processing is facilitated by computing parameters which are forwarded with traffic connection request messages that permit subsequent nodes to rapidly compute approximate EBRs. The approximate EBRs are adequately accurate to permit connections to be established with a high level of quality assurance. In accordance with the method, in order to reduce the computational effort, interpolation is used to derive approximate EBRs at subsequent links involved in the processing of traffic admission requests. In a method in accordance with the invention, a hyperbolic interpolation is used to yield a reliable approximation for EBR based on link capacity. Other types of interpolators may be used which could produce equally good results.

[0070] Fig. 14 shows a graph of a hyperbolic interpolator in accordance with the invention. Using the interpolator, when an edge node 22 receives a traffic admission request, it computes an EBR value for the connection using the capacity R_1 of one of the links which the node control element 28 found to have adequate free capacity for the admission connection request. Thereafter, the node control element computes values for Ω_∞, a, and b, and R using the following formulas:

$$\Omega_\infty = \gamma\theta ;$$

where

γ is the peak rate of the traffic stream requesting admission to the network; and
θ is the proportion of time that the source is active.

[0071] The value of b is computed using the formula:

$$b + \gamma = (R_1 - \gamma)\frac{\Omega_1 - \Omega_\infty}{\Omega - \Omega_1}$$

wherein

R_1 is a selected service rate, as indicated above;
$\hat{\Omega}$ is determined using the Gibbens-Hunt formula, which is well-known in the art; and
Ω_1 is determined using the extended Gibbens-Hunt method described in United States Patent Application No. 08/723,649 filed October 3, 1996 and entitled ADMISSION CONTROL IN AN ATM SWITCHING NODE, which is incorporated herein by reference. Although the above-referenced application is related only to ATM networks, the methods it teaches are also applicable to unfragmented variable-sized packets.

[0072] The value of a is computed using the formula:

$$a = (b + \gamma)(\hat{\Omega} - \Omega_\infty).$$

[0073] The values of Ω_∞, a and b are then passed in the service admission request message to other nodes involved in connection setup. On receipt of those values, the following formula is used compute the approximated EBR:

$$\Omega = \Omega_\infty + \frac{a}{R + b}$$

where:

Ω is an approximate EBR; and
R is the link capacity of the link selected in response to the traffic admission request.

[0074] Using the formula a node is enabled to compute an approximated EBR with much less computational effort than recomputing EBR using the extended Gibbens-Hunt method or some other similar method. This significantly speeds up traffic admission request processing and thereby enhances overall network efficiency. This method is particularly useful in establishing multicast connections since dozens or hundreds of EBRs may have to be computed in order to establish the multicast connections. Using this method significantly improves the efficiency of processing traffic admission requests for multicast connections.

[0075] The invention therefore provides a multi-class network which is capable of transmitting variable-size packets without packet deconstruction. The multi-class network operates efficiently with reliable quality assurance. Since admission control, connection routing and service-rate control are distributed at the node level of the network, control messaging overhead is minimized and network resources are available for transport functions.

[0076] In summary, the present application discloses A method and apparatus for control of a multi-class digital network are described. The network supports a plurality of digital services using dynamically-configured service bands to support various transport modes and qualities of service. Flexible controls with two degrees of freedom are used at the node level to dynamically configure the service bands. The first degree of freedom is realized using adaptive alternate routing to balance the traffic intensity across the network and, hence, increase the efficiency of the bands. The second degree of freedom is achieved through adaptive network partitioning to provide elasticity to the band structure. This permits band capacity to be adaptively reconfigured as connections are added and dropped. The switching nodes use rate controllers to divide the capacity of each link among the connection classes according to rules which ensure consistent service attributes for each band across the network. The rate controllers are adapted to support different transport modes and different packet sizes so that the deconstruction and reconstruction of packets at network edge devices is eliminated. The advantage is a flexible multi-class network which dynamically reconfigures to service traffic loads as service demands fluctuate.

[0077] Modifications of the above-described embodiments will no doubt become apparent to those skilled in the art. The scope of the invention is. therefore intended to be limited solely by the scope of the appended claims.

Claims

1. A method of controlling traffic admission and routing in a multi-class digital network serving variable-size packets, by computing an equivalent bit rate for each traffic admission request as received at a node in the network, determining whether a connection for the traffic admission request can be established through the network, and establishing the connection if a connection can be established, **CHARACTERIZED BY**:

   adding the equivalent bit rate for the traffic admission request to a current service rate allocation of a class

of traffic being served by an egress link through which the connection is established, to permit a service rate controller to control transmission of each class of traffic on the egress link.

2. A method as claimed in claim 1 wherein a determination of whether a connection for a traffic admission request can be established through the network includes steps of:

a) determining a class of service to which the traffic admission request belongs;
b) selecting a direct path to a destination for the traffic admission request if a direct path exists for the class of service to which the traffic admission request belongs, and the path has adequate bandwidth available to accommodate the traffic; else
c) selecting a routing method associated with the class of service;
d) examining, in a predetermined order, route sets associated with the destination to determine if a link exists with adequate available bandwidth to accommodate the traffic; and
e) if a link is found, attempting to establish a connection for the traffic on the link; else
f) rejecting the traffic admission request.

3. A method as claimed in claim 2 wherein the routing methods which may be associated with a class of service comprise:

a) shortest path hop-by-hop routing;
b) selective routing using a conservative scheme; and
c) selective routing using true-state information.

4. A method as claimed in claim 3 wherein the shortest path hop-by-hop routing comprises the steps of:

a) comparing in the predetermined order an occupancy state of links in a set of routes associated with the destination, and selecting a most vacant link in the set of routes where at least one link has adequate available bandwidth for the traffic admission request;
b) determining a number of hops in a shortest path to the destination for the traffic admission request given the selected link;
c) assigning a number of route selection credits to the traffic admission request, the number of route selection credits equalling the number of hops in the shortest available path, plus two;
d) forwarding a traffic admission request message to a node associated with an opposite end of the selected link;
e) subtracting 1 from the credits at the node associated with the opposite end of the selected link; and
f) repeating steps a), d) and e) without returning to an immediately preceding node until a connection is established to the destination, or the route selection credits are exhausted in which case the traffic admission request is rejected.

5. A method as claimed in claim 3 wherein the selective routing using a conservative scheme comprises the steps of:

a) selecting at most two links with adequate bandwidth to accommodate the admission request;
b) debiting an available capacity of the at most two links selected by the equivalent bit rate computed for the traffic admission request;
c) forwarding a traffic admission request message to a node associated with an opposite end of the at most two links;
d) repeating steps a)-c) until a connection for the traffic admission request is established or the traffic admission request is rejected; and
e) crediting an available capacity of any link debited which did not become a link in the connection after the connection is established or the traffic admission request is rejected.

6. A method as claimed in claim 3 wherein the selective routing using a true-state scheme comprises the steps of:

a) selecting at most two links with adequate bandwidth at an originating node to accommodate the traffic admission request;
b) declaring the at most two links unavailable to another route selection process of a same class of service until a connection to a destination is established or rejected;
c) forwarding a traffic admission request message to a node associated with an opposite end of the at most

two links;
d) determining if a next link in a route associated with each of the at most two links has adequate bandwidth to accommodate the traffic admission request and repeating steps b) and c) until a connection for the traffic admission request is established or the traffic admission request is rejected; and
e) if the connection is established, instructing from a controlling node all other nodes to debit an available capacity of links in the connection; else
f) declaring the at most two links available to other route selection processes.

7. A method as claimed in claim 3 where the three routing schemes coexist in the same network.

8. A method as claimed in any preceding claim wherein the combined service rate allocations for all classes of traffic being served by an egress link satisfies the condition:

$$0 < f_j = \frac{F_j}{R} < 1, j = 0,\ldots,K-1, \quad with \quad \sum_{j=0}^{K-1} f_j < 1$$

wherein:

K is the number of classes;
R is the link rate in bits per second; and
F_j is the required service allocation for class j in bits per second; and
f_j is the normalized service rate for class j.

9. A method as claimed in claim 1 wherein a link controller for a transport link in the multi-class digital network includes a service rate controller for controlling an egress of variable-sized packets on the transport link by performing the steps of:

a) storing a class service allocation computed by a node control element in response to a traffic load for the class at a specific point in time;
b) adding a value of the class service allocation to an adder memory after counting a predetermined number of clock signals;
c) comparing a value in the adder memory to a normalized packet size; and
d) visiting each comparator and writing a class number to a ready queue when the comparator determines that the value in the adder memory is greater than or equal to the normalized packet size to enable packets to be transferred onto the link when the class number reaches a head of the ready queue.

10. A method as claimed in claim 9 wherein the clock signals are produced by a clock that is local to the link controller.

11. A method as claimed in claim 10 wherein the clock produces clock signals at a rate which is equal to a time required to transfer a smallest packet accommodated by the link.

12. A method as claimed in claim 9 wherein the service rate controller is adapted to control an egress of variable-sized packets using different transport modes, each transport mode consisting of at least one transport protocol and traffic using different transport modes are assigned to different classes.

13. A method as claimed in claim 9 wherein the normalized packet size is representative of a time required to transfer a given packet.

14. A method as claimed in any preceding claim wherein computing the equivalent bit rate (EBR) for the traffic admission request comprises the steps of:

a) receiving from the traffic admission request source traffic descriptors which indicate the type and intensity of the traffic to be admitted to the network, or directly measuring the traffic to determine the type and intensity;
b) selecting a quality of service (QOS) for the traffic admission request; and
c) computing the EBR using the traffic descriptors, the QOS and the entire capacity of a link which will be used to transport the traffic if the admission request is accepted.

15. A method as claimed in claim 14 wherein the traffic descriptors include traffic peak rate, traffic mean rate and mean packet size.

16. A method as claimed in claim 14 or 15 further comprising the steps of:

a) computing at a node control element where a traffic admission request is received, the EBR and parameters which may be used in subsequent nodes to compute an approximated EBR, the approximated EBR being used to determine in subsequent nodes whether link capacity exists to accept the traffic admission request;
b) passing the parameters to an adjacent node control element;
c) computing at the adjacent node an approximated EBR using the parameters in a simplified interpolation formula that conserves computational effort; and
d) repeating steps b) and c) until a connection is established or the admission request is rejected.

17. A method as claimed in claim 16 wherein the parameters are computed using a hyperbolic approximation.

18. A method as claimed in claim 16 or 17 wherein three parameters are computed.

19. A method as claimed in claim 18 wherein the three parameters are Ω_∞, a and b and they are computed using the formulas:

$$\Omega_\infty = \gamma\theta ;$$

where

γ is the peak rate of the traffic stream requesting admission to the network; and
θ is proportion of time the source is active;
the value of b is computed by the formula:

$$b + \gamma = (R_1 - \gamma) \frac{\Omega_1 - \Omega_\infty}{\hat{\Omega} - \Omega_1}$$

wherein
R_1 is a selected service rate R;
$\hat{\Omega}$ is determined using the Gibbens-Hunt formula; and
Ω_1 is determined using an extended Gibbens-Hunt method;

the value of a is computed using the formula:

$$a = (b + \gamma)(\hat{\Omega} - \Omega_\infty).$$

20. A method as claimed in claims 16-19 wherein the traffic admission request is for traffic to be multicast.

21. A method as claimed in any preceding claim wherein the traffic in the network is classified by node control elements into three types, comprising:

a) connection-based traffic with a specified QOS;
b) connection-based traffic without a specified QOS; and
c) connectionless traffic without a QOS specification.

22. A method as claimed in claim 21 wherein each node control element monitors the traffic of type b) and allocates an appropriate service rate to that traffic if there is sufficient free capacity on a link used by the traffic until the capacity is required by traffic of type a).

23. A method as claimed in claim 21 or 22 wherein type c) traffic is transferred on a link only when there is no traffic of types a) or b) waiting to be transferred on the link.

24. Apparatus for distributed control of a multi-class network in which network processing is distributed between a network control element and network nodes to mimimize computational effort and maximize transmission efficiency, **CHARACTERIZED** in that:

the network control element receives traffic state and network state information form the nodesand uses the information to compute network traffic routing sets that are distributed to the nodes;
the network nodes include node control elements that control traffic admission, traffic routing and the computation of service rate allocations for classes of traffic served by egrees links at the node; and
link controllers for the links, the link controllers including service rate controllers adapted to control an egress of variable-sized packets on the link, at a guaranteed service rate.

25. Apparatus as claimed in claim 25 wherein the service rate controller comprises:

a) a sampling frequency memory for storing a class service allocation computed by a node control element in response to a traffic load for the class at a specific point in time;
b) an adder adapted to add a value of the class service allocation to an adder memory after counting a predetermined number of clock signals;
c) a comparator adapted to compare a value in the adder memory to a normalized packet size;
d) a selector for visiting each comparator and writing a class number to a ready queue when the comparator determines that the value in the adder memory is greater than or equal to the normalized packet size to enable packets to be transferred onto the link when the class number reaches a head of the ready queue.

26. Apparatus as claimed in claim 25 wherein a clock that produces the clock signals is a clock that is local to the link controller.

27. Appartus as claimed in claims 25 or 26 wherein the clock produces clock signals at a rate which is equal to a time required to transfer a smallest packet accommodated by the link.

28. Appatratuss as claimed in claims 25-27 wherein the service rate controller is adapted to control an egress of variable-sized packets using different transport modes, each transport mode consisting of at least one transport protocol and traffic using different transport modes are assigned to different classes.

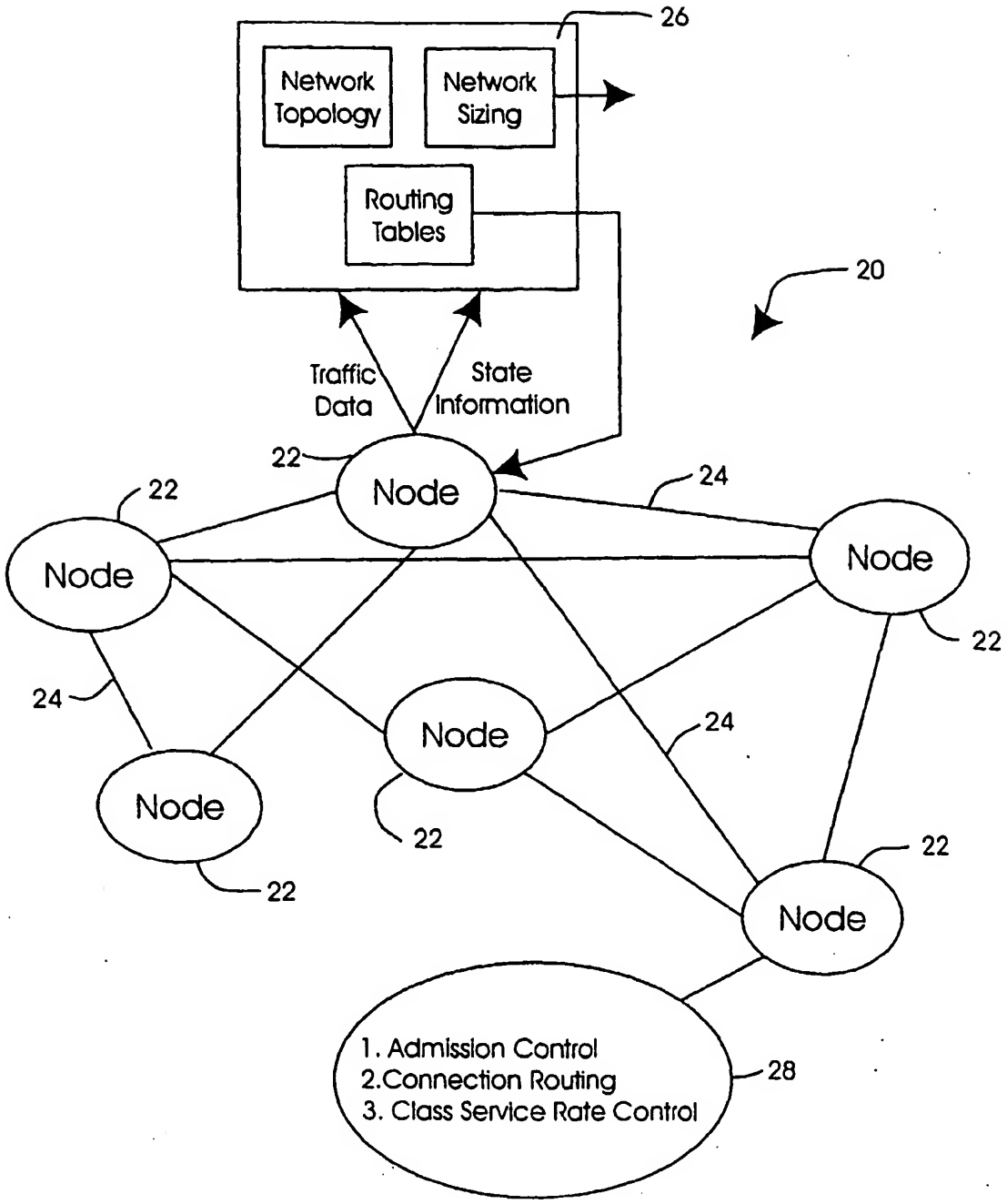

26
Network Topology
Network Sizing
Routing Tables
20
Traffic Data
State Information
22
Node
24
Node
Node
22
24
Node
24
Node
22
22
Node
22
1. Admission Control
2.Connection Routing
3. Class Service Rate Control
28

FIG. 1

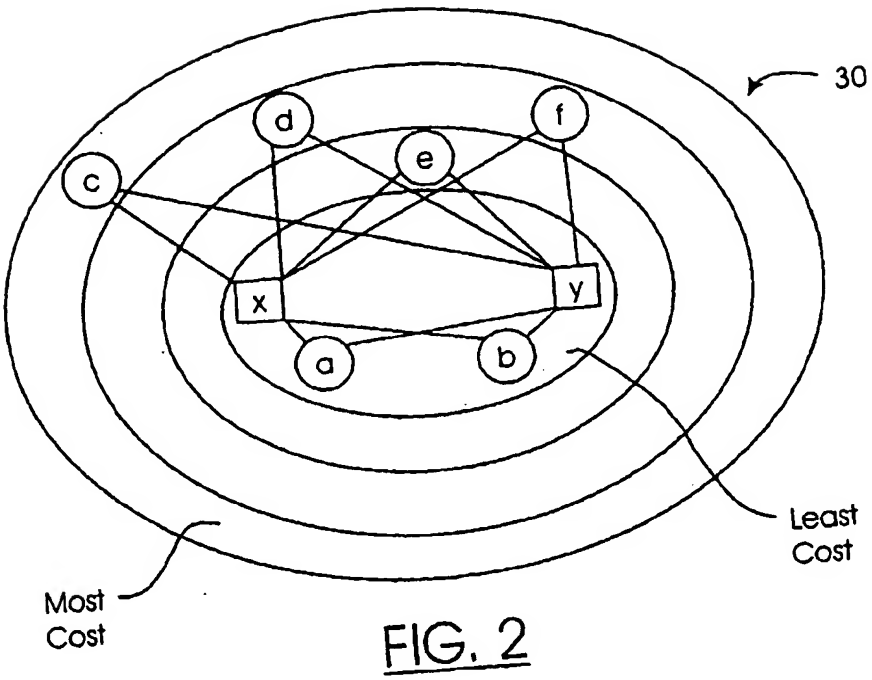


FIG. 2

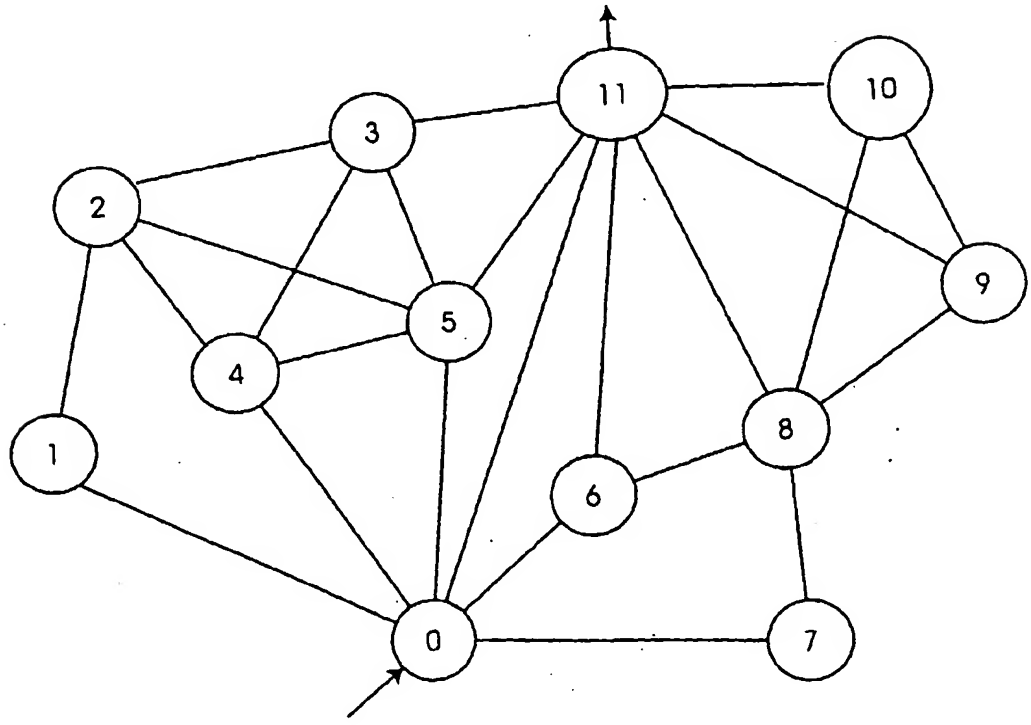


FIG. 3

Destination Node: 11

From Node	Next Node		
	Direct	First Set	Second Set
0	11	5 / 6	4 / 7
1		0	2
2		3 / 5	1 / 4
3	11	5	
4		3 / 5	2 / 0
5	11	3 / 0	4
6	11	8 / 0	
7		8 / 0	
8	11	10 / 9 / 6	7
9	11	8 / 10	
10	11	9 / 8	

FIG. 4

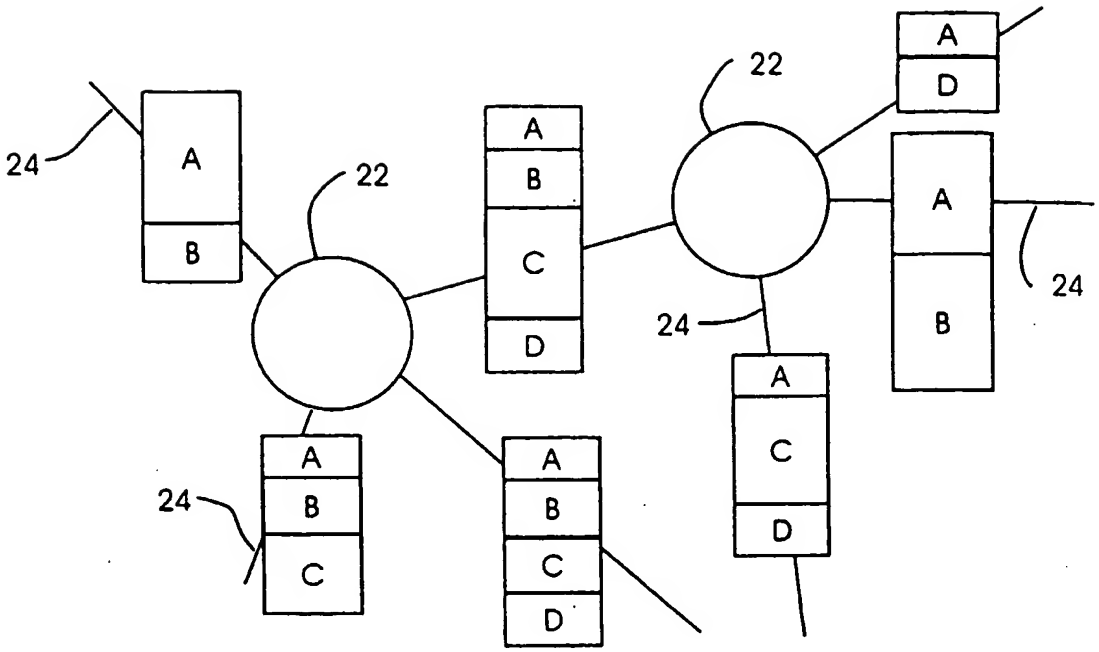

24
A
B
22
A
B
C
D
22
A
D
A
B
24
24
A
C
D
24
A
B
C
A
B
C
D

FIG. 5

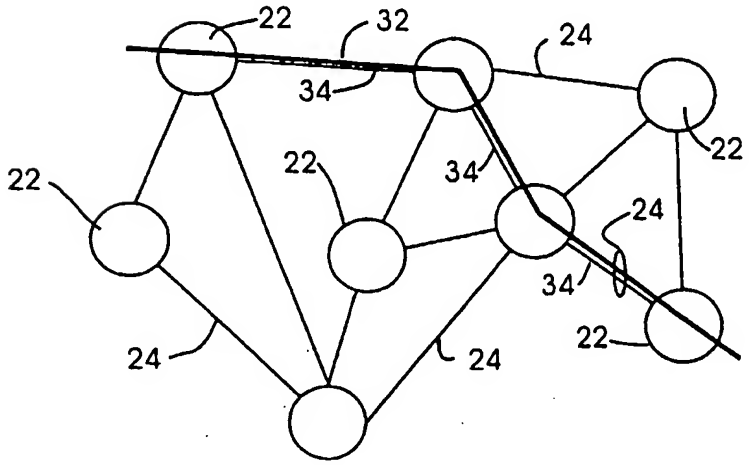

22
32
24
34
22
22
34
24
34
22
24
24
22

FIG.6

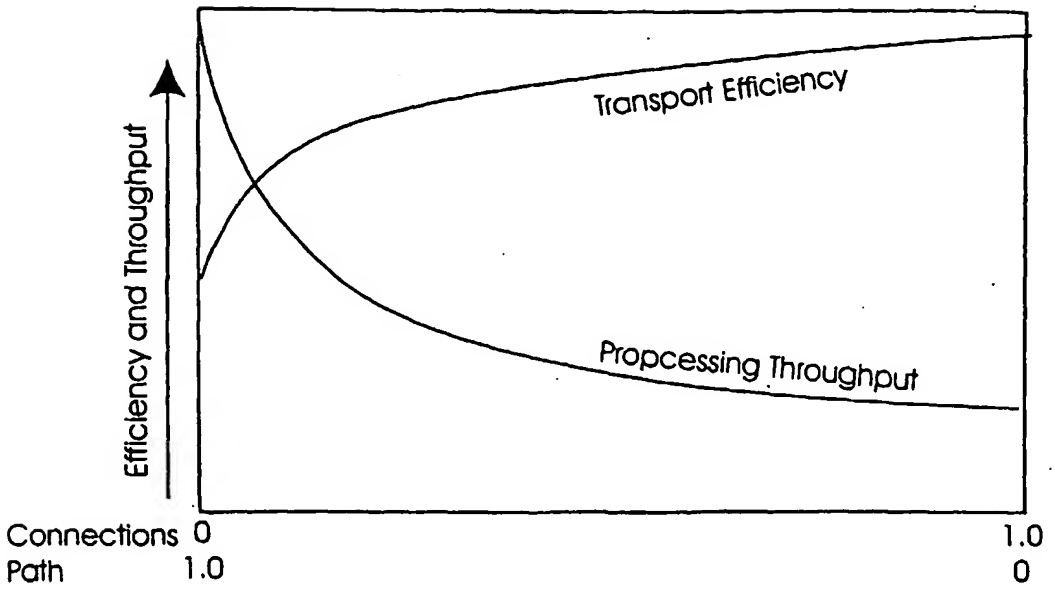

Transport Efficiency
Propcessing Throughput
Efficiency and Throughput
Connections 0
Path 1.0
1.0
0

FIG. 7

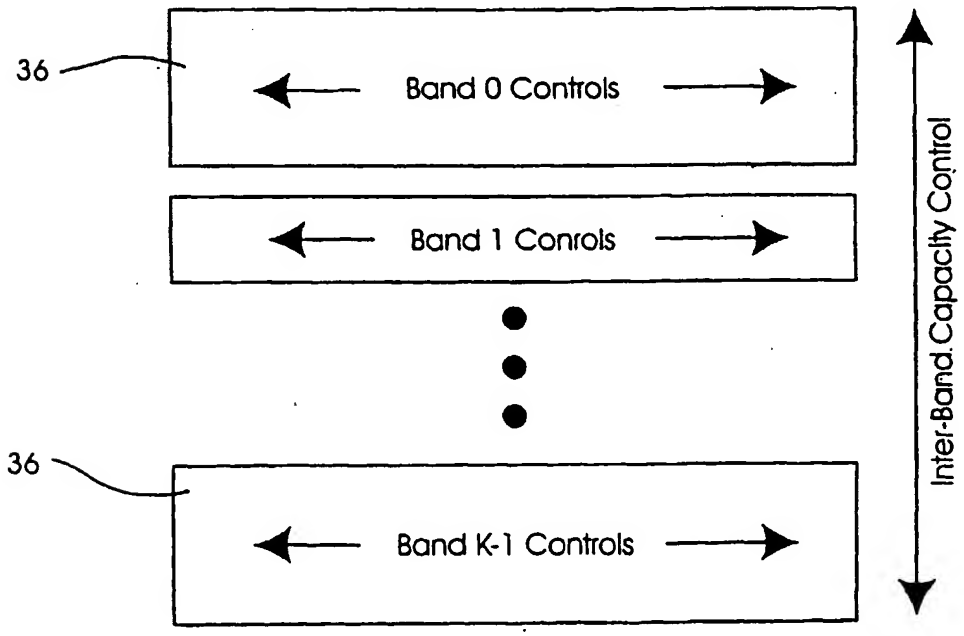

36
Band 0 Controls
Band 1 Conrols
36
Band K-1 Controls
Inter-Band Capacity Control

FIG. 8

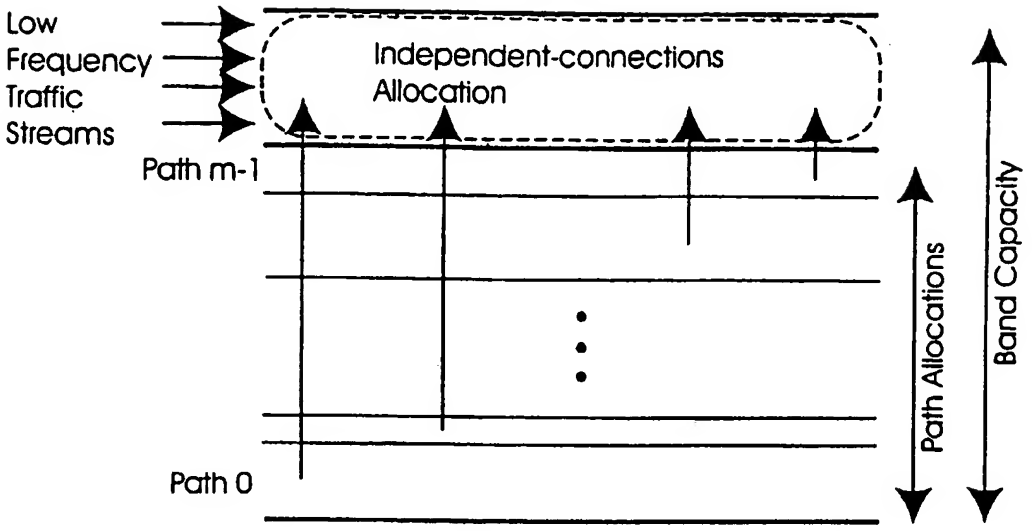

Low
Frequency
Traffic
Streams
Independent-connections
Allocation
Path m-1
Path 0
Path Allocations
Band Capacity

FIG. 9

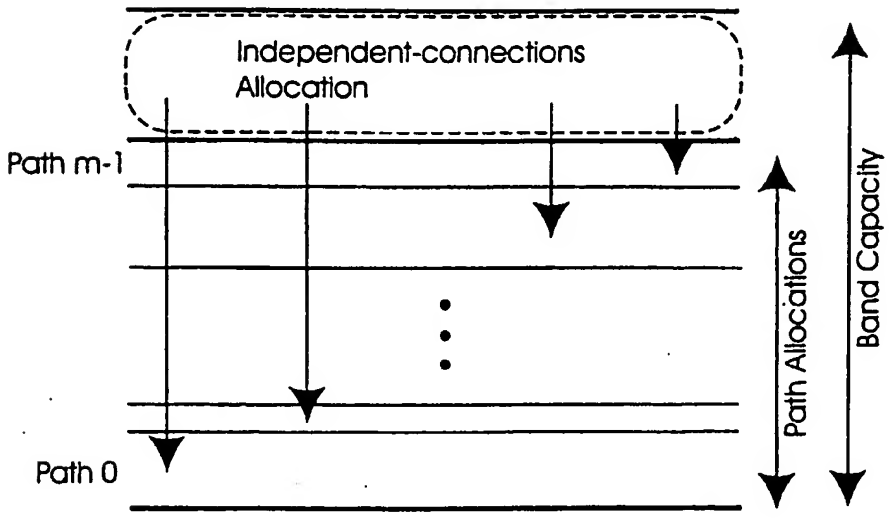

Independent-connections
Allocation
Path m-1
Path 0
Path Allocations
Band Capacity

FIG. 10a

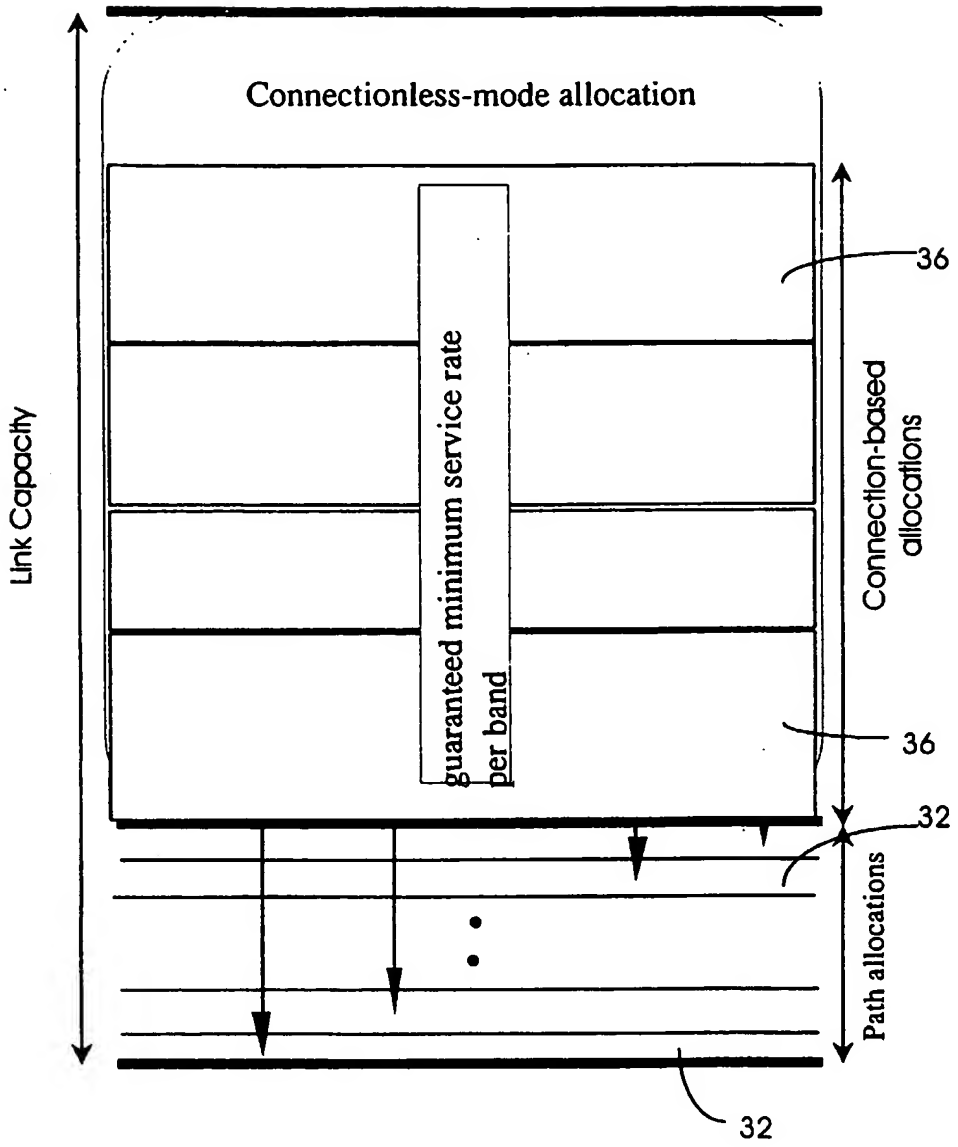

Connectionless-mode allocation
Link Capacity
guaranteed minimum service rate per band
Connection-based allocations
Path allocations
36
36
32
32

FIG. 10b

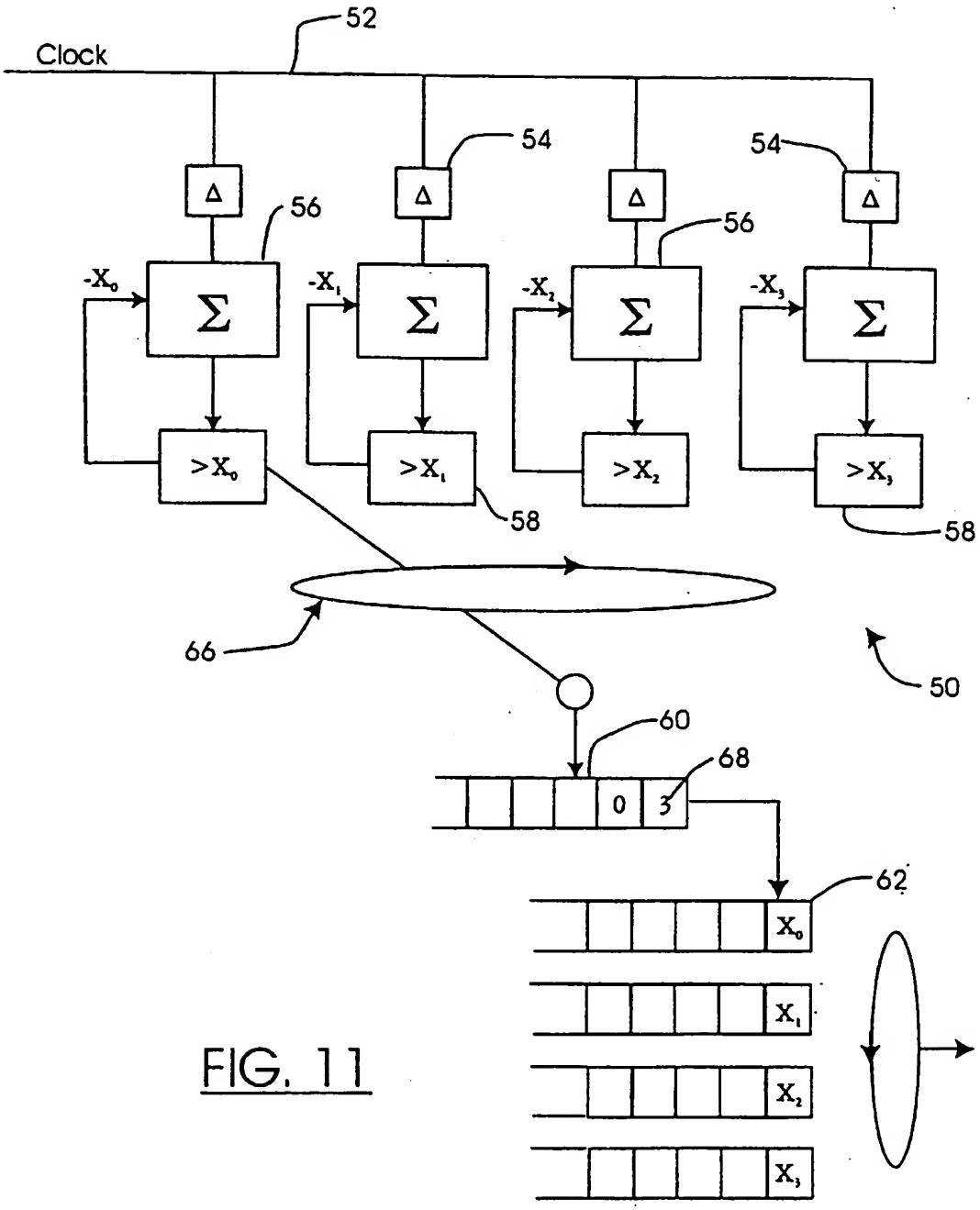
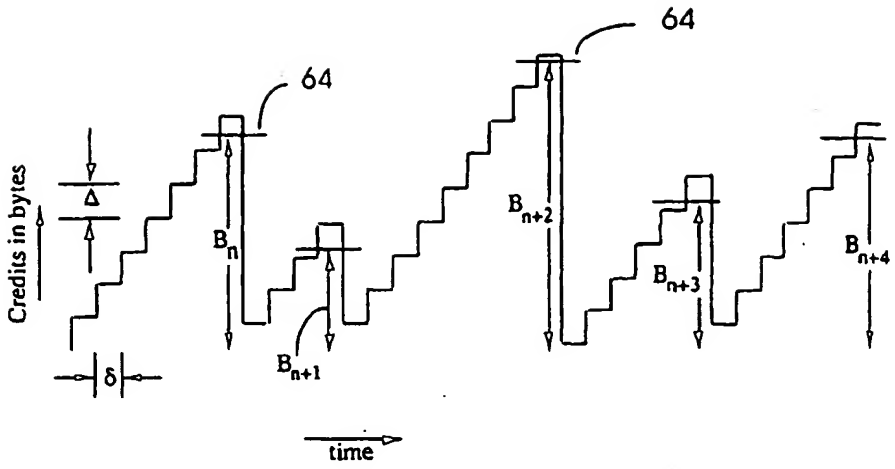

Clock
52
54
54
56
56
Δ
Σ
-X0
-X1
-X2
-X3
>X0
>X1
>X2
>X3
58
58
66
50
60
68
0
3
62
X0
X1
X2
X3

FIG. 11



Class Allocations

Δ = 0.29	0.21	0.30	0.20
0.29	0.21	0.30	0.20
0.58	0.42	0.60	0.40
0.87	0.63	0.90	0.60
1.16	0.84	1.20	0.80
1.45	1.05	1.50	1.00
1.74	1.26	1.80	1.20
2.03	1.47	0.30	1.40
2.32	1.68	0.60	1.60
2.61	0.29	0.90	1.80
2.90	0.50	1.20	2.00
0.39	0.71	1.50	2.20
0.68	0.92	1.80	2.40
0.97	1.13	2.10	0.28
1.26	1.34	2.40	0.48
1.55	1.55	0.50	0.68
1.84	0.36	0.80	0.88

δ

time

accumulated credit

1.60

1.80

64

56

2.80

64

2.32

64

1.40

56

2.20

FIG. 12

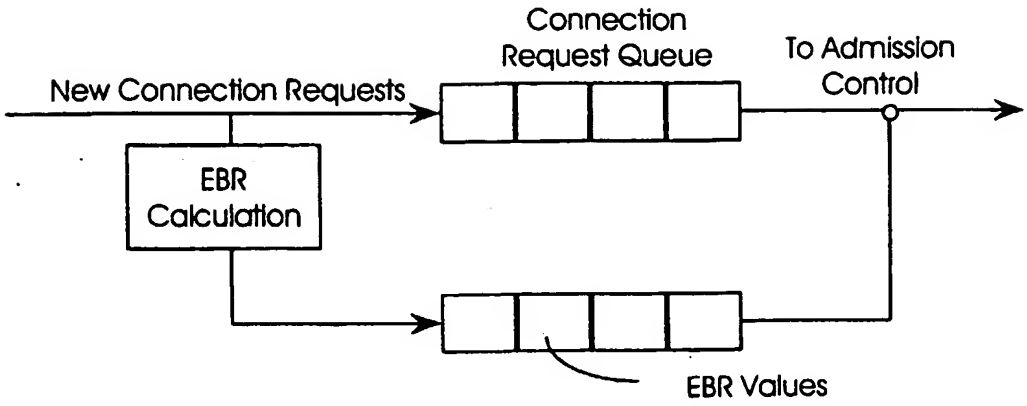

Connection
Request Queue
To Admission
Control
New Connection Requests
EBR
Calculation
EBR Values

FIG. 13

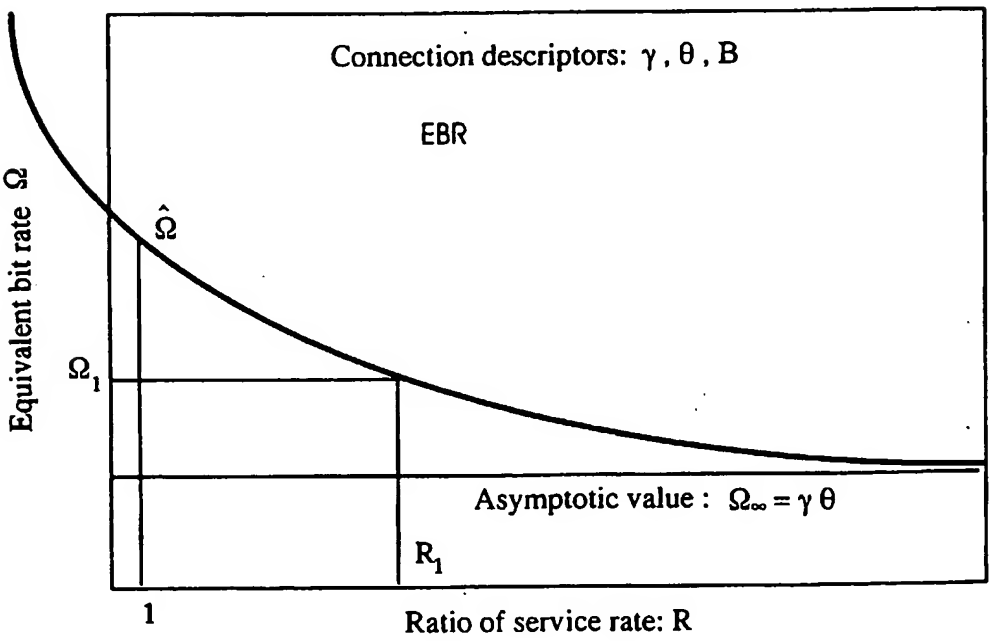

Connection descriptors: γ , θ , B
EBR
Equivalent bit rate Ω
Ω̂
Ω1
Asymptotic value : Ω∞ = γ θ
R1
1
Ratio of service rate: R

FIG. 14

(19) Europäisches Patentamt
Europ an Patent Office
Office européen des brevets

(11) EP 0 957 648 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
07.06.2000 Bulletin 2000/23

(43) Date of publication A2:
17.11.1999 Bulletin 1999/46

(21) Application number: **99303448.7**

(22) Date of filing: **30.04.1999**

(51) Int Cl.[7]: **H04Q 3/66**, H04Q 3/00

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **30.04.1998 CA 236427**

(71) Applicant: **NORTEL NETWORKS CORPORATION**
Montreal Quebec H2Y 3Y4 (CA)

(72) Inventors:
- **Beshai, Maged E.**
  Stittville, Ontario K2S 1p2 (CA)
- **Vickers, Richard**
  Kanata, Ontario K2K 1P2 (CA)

(74) Representative: **Bewley, Ewan Stuart**
Nortel Networks,
IP Law Group,
London Road
Harlow, Essex CM17 9NA (GB)

(54) Method and apparatus for distributed control of a multi-class network

(57) A method and apparatus for control of a multi-class digital network (20) are described. The network supports a plurality of digital services using dynamically-configured service bands (36) to support various transport modes and qualities of service. Flexible controls with two degrees of freedom are used at the node level to dynamically configure the service bands. The first degree of freedom is realized using adaptive alternate routing to balance the traffic intensity across the network and, hence, increase the efficiency of the bands. The second degree of freedom is achieved through adaptive network partitioning to provide elasticity to the band structure. This permits band capacity to be adaptively reconfigured as connections are added and dropped. The switching nodes (22) use rate controllers (50) to divide the capacity of each link (24) among the connection classes according to rules which ensure consistent service attributes for each band across the network. The rate controllers are adapted to support different transport modes and different packet sizes so that the deconstruction and reconstruction of packets at network edge devices is eliminated. The advantage is a flexible multi-class network which dynamically reconfigures to service traffic loads as service demands fluctuate.

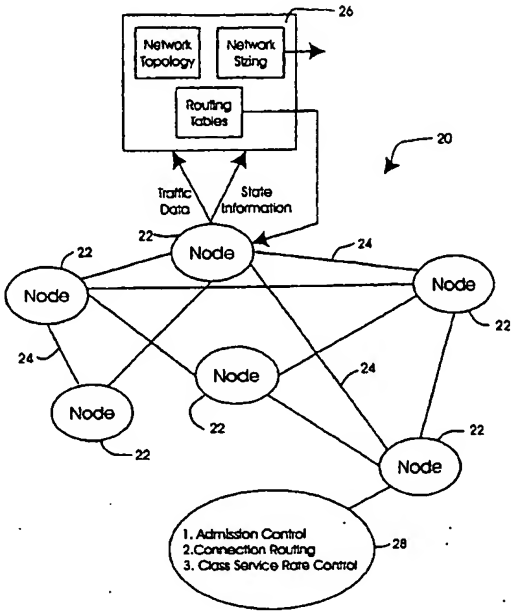


FIG. 1

EP 0 957 648 A3

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 30 3448

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
A	ERIKSSON A: "REAL-TIME SERVICES OVER THE INTERNET" ISS. WORLD TELECOMMUNICATIONS CONGRESS. (INTERNATIONAL SWITCHING SYMPOSIUM),CA,TORONTO, PINNACLE GROUP, 1997, pages 173-179, XP000704466 * page 175, line 16 - line 55 * ---	1,24	H04Q3/66 H04Q3/00 H04L12/56
A	YAN J ET AL: "TRAFFIC CONSIDERATIONS IN THE SYNTHESIS OF AN ATM-BASED NETWORK" ITC SPECIALISTS SEMINAR,NL,LEIDSCHENDAM, KPN RESEARCH, vol. SEMINAR 9, 1995, pages 27-44, XP000683144 * page 29, line 18 - page 41, line 11 * ---	1-3,14, 15,21-24	
A	US 5 629 930 A (BESHAI MAGED E ET AL) 13 May 1997 (1997-05-13) * column 5, line 53 - column 6, line 51 * -----	5,6	

TECHNICAL FIELDS SEARCHED (Int.Cl.6)

H04Q
H04L

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	18 April 2000	Ströbeck, A

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 99 30 3448

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

18-04-2000

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5629930 A	13-05-1997	WO 9716907 A	09-05-1997
		EP 0858704 A	19-08-1998
		JP 10512426 T	24-11-1998

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82